

The Battle of Neighborhoods - Chennai

Applied Data Science Capstone by IBM on Coursera

(ROHITH S P)

1. INTRODUCTION: BUSINESS PROBLEM

This project deals with discussing the neighborhoods of **Chennai, The Detroit of India**. This project would specifically help Business people planning to start Restaurants, Hotels, etc. in Chennai, Tamil Nadu, India.

The **Foursquare API** is used to access the venues in the neighborhoods. Since, it returns less venues in the neighborhoods, we would be analyzing areas for which countable number of venues are obtained. Then they are clustered based on their venues using Data Science Techniques. Here the **k-means clustering algorithm** is used to achieve the task. The optimal number of clusters can be obtained using **silhouette score** metrics.

Folium visualization library can be used to visualize the clusters superimposed on the map of Chennai city. These clusters can be analyzed to help small scale business owners select a suitable location for their need such as Hotels, Shopping Malls, Restaurants or even specifically Indian restaurants or Coffee shops.

2. DATA REQUIREMENTS

Chennai has multiple neighborhoods. The chennaiiq.com website has a dataset which has the list of locations in Chennai along with their Latitude and Longitude. In order to obtain the venue details in each neighborhood Foursquare API is used.

1. https://chennaiiq.com/chennai/latitude_longitude_areas.asp
2. <https://foursquare.com/>

There is a total of 105 neighborhoods. But the Latitude and Longitude data obtained are in Degrees Minute Seconds format which needs to be converted to Decimal Degrees Format. The following data are obtained from the Foursquare API,

- Venue
- Venue Latitude
- Venue Longitude
- Venue Category data

A total of 1130 venues data have been obtained from Foursquare. The resultant venues dataset, (shown in Fig. 1) is used for the analysis process.

| <pre>print(chennai_venues.shape) chennai_venues.head()</pre> | | | | | | | |
|---|-----------------|-----------------------|------------------------|-----------------------------------|----------------|-----------------|---------------------------|
| (1130, 7) | | | | | | | |
| | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
| 0 | Adyar Bus Debot | 12.997222 | 80.256944 | Zaitoon Restaurant | 12.996861 | 80.256178 | Middle Eastern Restaurant |
| 1 | Adyar Bus Debot | 12.997222 | 80.256944 | Kuttanadu Restaurant | 12.997010 | 80.257799 | Asian Restaurant |
| 2 | Adyar Bus Debot | 12.997222 | 80.256944 | Zha Cafe | 12.999730 | 80.254806 | Café |
| 3 | Adyar Bus Debot | 12.997222 | 80.256944 | Adyar Ananda Bhavan, Besant Nagar | 12.996678 | 80.258275 | Fast Food Restaurant |
| 4 | Adyar Bus Debot | 12.997222 | 80.256944 | Kovai Pazhamudir Nilayam | 12.996522 | 80.259776 | Fruit & Vegetable Store |
| A total of 1130 venues were obtained. Now lets check the number of venues returned per neighbourhood. | | | | | | | |

Fig. 1 Chennai Venues Dataset